

Precoding for the Sparsely Spread MC-CDMA Downlink with Discrete-Alphabet Inputs

Min Li, *Member, IEEE*, Chunshan Liu, *Member, IEEE*, and Stephen V. Hanly, *Senior Member, IEEE*

Abstract—Sparse signatures have been proposed for the CDMA uplink to reduce multi-user detection complexity, but they have not yet been fully exploited for its downlink counterpart. In this work, we propose a Multi-Carrier CDMA (MC-CDMA) downlink communication, where regular sparse signatures are deployed in the frequency domain. Taking the symbol detection point of view, we formulate a problem appropriate for the downlink with discrete alphabets as inputs. The solution to the problem provides a power-efficient precoding algorithm for the base station, subject to minimum symbol error probability (SEP) requirements at the mobile stations. In the algorithm, signature sparsity is shown to be crucial for reducing precoding complexity. Numerical results confirm system-load-dependent power reduction gain from the proposed precoding over the zero-forcing precoding and the regularized zero-forcing precoding with optimized regularization parameter under the same SEP targets. For a fixed system load, it is also demonstrated that sparse MC-CDMA with a proper choice of sparsity level attains almost the same power efficiency and link throughput as that of dense MC-CDMA yet with reduced precoding complexity, thanks to the sparse signatures.

Index Terms—CDMA, discrete alphabets, MC-CDMA, power efficiency, precoding, sparse signature, symbol error probability.

I. INTRODUCTION

A. Motivations and Contributions

Multi-Carrier Code Division Multiple Access (MC-CDMA) is a multi-access scheme based on the Orthogonal Frequency Division Multiplexing (OFDM) method. Since its invention, MC-CDMA has attracted broad interest, see, e.g., [1]–[4] and the references therein. MC-CDMA naturally integrates CDMA's flexible multiuser access with interference suppression capability and the advantages of multicarrier OFDM, including robustness against frequency-selective fading. Therefore, it has the potential to be one of the candidates to support massive access and provide reliable data communication and better coverage for future-generation wireless systems.

As in all CDMA systems, MC-CDMA may experience severe multi-access interference due to the loss of user orthogonality, which may occur, particularly, in frequency-selective

channel environments. For such systems, optimal detection entails an exponential number of hypothesis testings about data symbols of all users and thus could be computationally demanding, time-consuming and even infeasible in a large system with conventional dense signature design. To circumvent the complexity issue, sparse signatures, whose fraction of non-zero entries is small, have been introduced and exploited for the CDMA uplink multi-user detection [5]–[10]. In particular, the belief-propagation algorithm has been proposed for such a system, an algorithm that has a natural implementation using parallel computation units, and one that is fast and provably optimal for different ensembles of sparse signatures in the large system limits [5]–[8]. Inspired by the belief-propagation algorithm, references [9] and [10] have developed reduced-complexity soft-in soft-out (SISO) and Turbo iterative multiuser detection algorithms for the sparse CDMA uplink and a sparse-signature OFDM uplink, respectively.

In this work, we formulate a different problem, appropriate for the MC-CDMA downlink counterpart, from the symbol detection point of view. The solution to the problem provides a power-efficient precoding algorithm for the base station (BS). The precoding is implemented at the BS, allowing the mobile stations (MSs) to use simple conventional single-user matched filters and standard single-user symbol detection. This hence simplifies the implementation of the receiver, as compared to the conventional MC-CDMA downlink transceiver design, where multi-user interference is normally mitigated by a frequency-domain equalization at the receiver [2]. Moreover, the proposed algorithm optimizes the signals transmitted on different subcarriers so that they can be constructively combined at each receiver, leading to a power-efficient precoding. The use of random sparse signatures was suggested for the MC-CDMA downlink in [11] to allow low-complexity iterative multiuser detection at each MS. But here we show that sparsity can be exploited to reduce the complexity of the precoding, as compared to MC-CDMA with dense signatures. In addition, using sparse signatures simplifies channel measurement, since each MS only needs to estimate channels for a small number of subcarriers it occupies.

The contributions of this paper are summarized as follows:

- We introduce the MC-CDMA downlink communication with regular sparse signatures, where each MS has access to an equal number of subcarriers and each subcarrier has (roughly) equal load. We also consider a bipartite graph representation for the system studied and use it to facilitate algorithm design and complexity analysis.
- Assume that data symbols intended for MSs are drawn from discrete alphabets. We take the symbol detection

Copyright (c) 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

This work was presented in part in Proceedings of the IEEE International Conference on Communications (ICC), Sydney, Australia, June 2014. This research was supported in part by the Australian Research Council under grant DP130101760, and by the CSIRO Macquarie University Chair in Wireless Communications. This Chair has been established with funding provided by the Science and Industry Endowment Fund.

The authors are with the Department of Engineering, Macquarie University, Macquarie Park, NSW 2113, Australia (e-mail: {min.li, chunshan.liu, stephen.hanly}@mq.edu.au).

point of view and introduce the minimum Symbol Error Probability (SEP) as a Quality of Service (QoS) metric for each MS. We formulate the precoding problem as a transmit power optimization problem subject to minimum SEP requirements at MSs. We translate the SEP targets into constraint regions on noiseless received signal components at MSs and characterize them via a conservative approximation to make the problem tractable. Detailed formulation procedures are provided for systems with both standard 4/16-QAM constellations and Tomlinson-Harashima replica points.

- We develop a precoding algorithm that accommodates parallel computation units via the dual-decomposition theory. Aided by the graph representation of the system, the complexity of the algorithm is characterized in terms of the number of message passings between computation units and the number of additions and multiplications for precoding calculation. Signature sparsity is shown to play a vital role in reducing precoding complexity.
- We demonstrate that the proposed optimized precoding generally outperforms the conventional zero-forcing (ZF) and the optimized Regularized ZF (RZF) precodings in terms of power efficiency under the same SEP requirements. The exact gain depends on the system load and it is very significant for a fully loaded system. We also demonstrate that, for a fixed system load, sparse MC-CDMA, e.g., with a proper relatively small number of subcarriers allocated to each MS, attains almost the same power efficiency and link throughput as dense MC-CDMA under our proposed precoding scheme. This important observation, in conjunction with the fact that sparsity reduces precoding complexity, promotes the practicality of sparse MC-CDMA.

B. Other Related Work

Precoding is a relatively mature concept in multi-antenna communication systems, enabling multiuser multiplexing in the spatial domain, see, e.g., [12]–[20]. It has to balance between the two conflicting interests of maximizing the useful signal power at the intended user and minimizing interference leakage towards non-intended users. The same concept can be applied to other systems such as direct-sequence CDMA or MC-CDMA, where multiplexing takes place in the time- or frequency-domain, and multi-access interference, if it arises, has to be dealt with [21]–[25].

Existing precoding techniques can be divided into two categories, linear precoding and non-linear precoding, where both require channel state information while the latter requires additional symbol-based processing. Within the first category, matched filtering, ZF and RZF [12] are three commonly known precodings that maintain different levels of balance between the two conflicting goals. Other power-aware linear precodings of general form have also been proposed in the literature subject to different QoS metrics, such as signal-to-noise-plus-interference ratio [20].

Compared with linear precoding, non-linear precoding may offer higher power efficiency and transmission rate, but the

gain comes at the cost of incorporating more sophisticated signal processing [13], [15]–[17], [23]–[25]. In the multi-antenna broadcasting setup, capacity-achieving non-linear Dirty-Paper Coding (DPC) entails a successive pre-cancellation of known intra-user interference at the BS. The encoding of data relies on codewords of infinite length and involves a high-dimensional sphere-search algorithm, which renders DPC unattractive in practical systems. Tomlinson-Harashima precoding (THP) is a simplified version of DPC, where the codebook is comprised of periodic extension of standard constellations (replica points) in the two-dimensional space and a transmit modulo-operation is introduced in the interference pre-cancellation process in order to reduce transmit power. Built on ZF or RZF, reference [13] generalizes the single-user-based symbol extension idea of THP and introduces a joint perturbation of user symbol vector to further reduce transmit power.

The optimized precoding proposed in this work belongs to the second category. As in existing works, power consumption is one of the primary concerns in our optimized precoding. However, the optimization criterion, minimum SEP constraint, has not been considered before, except our own works [26]–[28] in the MIMO (or distributed MIMO) setup. This criterion appears to be natural when we consider a system with discrete alphabets as inputs. In addition, in our formulation, we fix the information-bearing alphabets at the BS but allow a certain relaxation of received signals at MSs through precoding, as long as they reside in detection-favorable regions and the SEP targets are met. This is distinguished from related works [16], [18], [19], [25], where non-linear relaxation of input alphabets has been adopted. In [16], [18], the relaxation is required to maintain the minimum signalling distance, while in [19], [25], the relaxation is to ensure the corresponding symbol-energy-to-noise ratio is above a certain threshold [19], [25]. However, in all these works, no explicit SEP targets are imposed at MSs.

Notation: Boldface uppercase and lowercase letters denote matrices and vectors, respectively, e.g., \mathbf{A} is a matrix and \mathbf{a} is a vector; \mathbf{I}_N is an $N \times N$ identity matrix; for integers $i \leq j$, $[i : j] = \{i, i+1, \dots, j\}$, is the discrete interval between i and j , and $\mathbf{a}_{[i:j]} = \{a_i, \dots, a_j\}$, is the collection of $[i : j]$ th components of vector \mathbf{a} ; $(\cdot)^T$ denotes the matrix transpose, while $(\cdot)^\dagger$ denotes the conjugate transpose; notation $E[X]$ denotes the expectation operation on random variable X , and $\lfloor x \rfloor$ denotes a floor function of real number x ; $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary part of a complex number, respectively; finally, $\|\mathbf{A}\|_p$ is the standard l_p norm of \mathbf{A} .

II. SYSTEM MODEL

A. Signaling Model

We consider a downlink communication, where a single-antenna BS is simultaneously serving K single-antenna MSs via MC-CDMA. Specifically, data symbols intended for MSs are all drawn from discrete alphabet sets, e.g., M -QAM constellations, common in practical deployments. The downlink communication takes place over a set of N orthogonal subcarriers where we assume $N \geq K$ and thus the load $\alpha = K/N \in (0, 1]$. In the conventional MC-CDMA, the

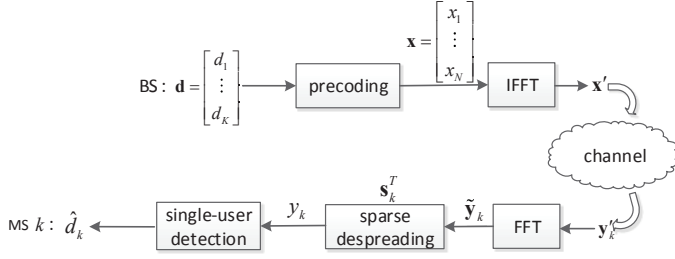


Fig. 1. The transceiver architecture for the sparse MC-CDMA downlink.

data symbol intended for a MS is transmitted over all parallel subcarriers where each is encoded with a binary phase-offset [1]. Here, however, information associated with each MS is assumed to be spread onto only a small subset of the available subcarriers, which leads to the sparse MC-CDMA as originally studied by [5]–[8] for the uplink.

Let $\mathbf{s}_k = \frac{1}{\sqrt{L_k}} [\tilde{s}_{k,1}, \tilde{s}_{k,2}, \dots, \tilde{s}_{k,N}]^T$ be the signature for MS k , where normalization factor L_k corresponds to the total number of subcarriers allocated to MS k . In the signature, components $\tilde{s}_{k,n}$ are i.i.d. drawn from a distribution P_S with zero-mean and unit-variance if MS k has access to subcarrier n , and $\tilde{s}_{k,n} = 0$ otherwise. The collection of all signatures corresponds to a sparse signature matrix $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_K]$, which is perfectly known at the BS.

The transceiver architecture for the downlink transmission is depicted in Fig. 1 and is elaborated as follows. Let $\mathbf{d} = [d_1, \dots, d_K]^T$ be the transmitted data symbol vector, where component d_k denotes the symbol intended for MS k and is drawn from a discrete finite-alphabet set \mathcal{D}_k . The transmission takes place by first forming appropriate frequency-domain signals and then converting them into time-domain signals by the inverse fast Fourier transform (IFFT) at the BS. Specifically, symbol vector \mathbf{d} is passed through a precoder and mapped to coded vector $\mathbf{x} \in \mathbb{C}^N$. An IFFT is then applied over the coded vector in order to generate time-domain signal vector $\mathbf{x}' \in \mathbb{C}^N$ that is subsequently transmitted over the wireless channel.

Upon observing channel output, each MS k first performs an FFT and produces the frequency domain signal $\tilde{\mathbf{y}}_k$ as

$$\tilde{\mathbf{y}}_k = \tilde{\mathbf{h}}_k \circ \mathbf{x} + \tilde{\mathbf{z}}_k, \quad (1)$$

where notation “ \circ ” denotes the Hadamard product; vector $\tilde{\mathbf{h}}_k = [\tilde{h}_{k,1}, \tilde{h}_{k,2}, \dots, \tilde{h}_{k,N}]^T$ is the collection of frequency-domain channel gains from BS to MS k ; vector $\tilde{\mathbf{z}}_k$ is a circularly symmetric complex Gaussian noise with $\mathbb{E}[\tilde{\mathbf{z}}_k \tilde{\mathbf{z}}_k^\dagger] = N_0 \mathbf{I}_N$. Despreading is then performed at MS k based on its own signature \mathbf{s}_k , followed by a simple single-user detection. The corresponding output signal y_k after despreading is given by

$$y_k = \sum_{n=1}^N \underbrace{(s_{k,n} \tilde{h}_{k,n})}_{h_{k,n}} x_n + \underbrace{\mathbf{s}_k^T \tilde{\mathbf{z}}_k}_{z_k} = \sum_{n=1}^N h_{k,n} x_n + z_k, \quad (2)$$

where $s_{k,n} = \tilde{s}_{k,n} / \sqrt{L_k}$ is the n th component of \mathbf{s}_k , and each equivalent channel noise z_k is circularly symmetric complex Gaussian with zero mean and variance N_0 . Collecting all

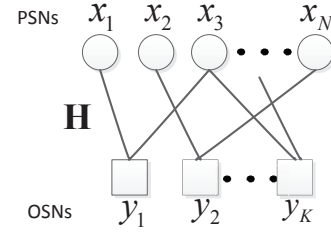


Fig. 2. Bipartite graph representation of the sparse MC-CDMA downlink.

outputs at MSs, we obtain the equivalent system input-output relationship as

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_K \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} h_{1,1} & \cdots & \cdots & h_{1,N} \\ h_{2,1} & \cdots & \cdots & h_{2,N} \\ \vdots & \vdots & \vdots & \vdots \\ h_{K,1} & \cdots & \cdots & h_{K,N} \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}}_{\mathbf{x}} + \underbrace{\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_K \end{bmatrix}}_{\mathbf{z}}. \quad (3)$$

It is straightforward to observe that in matrix \mathbf{H} , $h_{k,n} = 0$ as long as $s_{k,n} = 0$, and thus row \mathbf{h}_k maintains the same level of sparsity as the corresponding signature \mathbf{s}_k . This also means the BS only needs to know the small number of $h_{k,n}$ s for which $s_{k,n} \neq 0$ for the purpose of precoding.

B. Graph Representation

Given the matrix \mathbf{H} from (3), we can alternatively construct a bipartite graph representation of the sparse MC-CDMA system. Assume that each symbol x_n in the graph is represented by a precoded symbol node (PSN), and each output y_k is represented by a output symbol node (OSN). PSNs will be drawn as circles and OSNs will be drawn as squares in the graph. PSN x_n is connected with OSN y_k only if $s_{k,n} \neq 0$ and $h_{k,n}$ is the weight associated with the edge. Fig. 2 depicts an instance of the graph \mathcal{G} for $L_k = 2$, where each OSN is connected with two PSNs. We use $\mathcal{I}(x_n)$ to denote the collection of OSNs connected to x_n and define the node degree of x_n as the cardinality $|\mathcal{I}(x_n)|$. Similarly, we use $\mathcal{I}(y_k)$ to denote the collection of PSNs connected to y_k and define the node degree of y_k as the cardinality $|\mathcal{I}(y_k)|$. This graph representation introduced will facilitate the description of the precoding algorithm and the corresponding complexity analysis in Section V.

C. Sparse Signature Ensemble

In the signature matrix, we assume that the non-zero elements $\{\tilde{s}_{k,n}\}$ are i.i.d. drawn from a uniform distribution on $\{+1, -1\}$. It is observed that generating non-zero elements according to other distributions, e.g., Gaussian distribution, has little impact on the averaged system performance. Hence, we stick to the binary uniform distribution, which leads to a binary phase-offset for the symbol transmitted as in [6], [7].

Depending on the number of subcarriers allocated across MSs and the load per subcarrier, we have three common signature ensembles suggested for the uplink [7]: i) irregular ensemble, where Poisson-distributed number of subcarriers

are allocated across MSs and the load per subcarrier is also Poisson-distributed; *ii*) semi-regular ensemble, where each MS is allocated a fixed positive integer number of subcarriers and the load per subcarrier is Poisson-distributed; and *iii*) regular ensemble, where the number of subcarriers allocated for each MS and the loading per subcarrier both take fixed positive integer values. In particular, [7] advocates the regular ensemble as it amongst others prevents the systematic inefficiency due to leaving some subcarriers unoccupied by any of MSs.

In this work, we follow [7] and deploy the regular-type ensemble to ensure that the system enjoys full utilization of resources and provides user fairness. Specifically, when the system is fully-loaded ($\alpha = 1$), a perfectly regular signature matrix is randomly generated in the sense that each MS is allocated L subcarriers, and each subcarrier is accessed by exactly L MSs. When the system is under-loaded ($\alpha < 1$), a nearly regular signature matrix is randomly generated such that each MS is allocated L subcarriers, and each subcarrier is of almost equal load, namely, accessed by either $\lfloor \alpha L \rfloor$ or $(\lfloor \alpha L \rfloor + 1)$ MSs.

III. OPTIMIZED PRECODING WITH SEP TARGETS

For notational convenience, we define $\bar{y}_k = \mathbf{h}_k^T \mathbf{x}$ as the noiseless received component at MS k (\mathbf{h}_k^T is the k th row of matrix \mathbf{H}) with real part $\bar{y}_k^{(r)} = \Re\{\bar{y}_k\}$ and imaginary part $\bar{y}_k^{(i)} = \Im\{\bar{y}_k\}$; similarly, denote the real and imaginary parts of data symbol and noise as $d_k^{(r)} = \Re\{d_k\}$ and $d_k^{(i)} = \Im\{d_k\}$, and $z_k^{(r)} = \Re\{z_k\}$ and $z_k^{(i)} = \Im\{z_k\}$, respectively. In addition, define $\sigma^2 = N_0/2$ as the fixed noise variance per signal dimension.

Assuming that all data symbols transmitted are selected from discrete alphabets, we take a symbol detection point of view and impose minimum Symbol Error Probabilities (SEPs) as the user QoS constraints. Specifically, let $\mathcal{A}(d_k)$ denote the decision region associated with data symbol d_k intended for MS k and Pe_k denote the SEP target. Detection error happens when the output signal y_k lies outside decision region $\mathcal{A}(d_k)$. According to the SEP requirement, the probability of error events should be no greater than the target, i.e.,

$$\Pr(y_k = (\bar{y}_k + z_k) \notin \mathcal{A}(d_k)) \leq Pe_k. \quad (4)$$

A question we then ask is: How do we design a precoder that efficiently maps a symbol vector \mathbf{d} into a precoded vector \mathbf{x} such that the SEP requirements at MSs can be met?

Following the conventional zero-forcing (ZF) approach, one could form \mathbf{x} according to

$$\mathbf{x} = \mathbf{H}^\dagger (\mathbf{H}\mathbf{H}^\dagger)^{-1} \mathbf{d}, \quad (5)$$

which inverts the channel matrix and forces noiseless component \bar{y}_k to sit exactly at the constellation point d_k . In order to meet a given SEP target, Pe_k , data symbol d_k has to be chosen from a discrete alphabet set whose minimum distance between any two neighboring points (denoted by d_{\min}) is above a certain threshold.

For instance, consider a system with M -QAM modulation whose standard constellation is represented by

$$\mathcal{D}_S = \left\{ a_R + ja_I \mid a_R, a_I \in \{\pm 1, \pm 3, \dots, \pm(\sqrt{M} - 1)\} \right\} \quad (6)$$

with $d_{\min} = 2$. We need to scale the constellation points (increasing the minimum distance d_{\min} , but also the transmit power) in order to meet the SEP target, Pe_k . Considering the 4-QAM constellation (see Fig. 3), the scaling factor, β_k , to use for MS k , must satisfy

$$\frac{1}{\sqrt{2\pi}\sigma} \int_{-\beta_k}^{+\infty} e^{-\frac{(z_k^{(r)})^2}{2\sigma^2}} dz_k^{(r)} \times \frac{1}{\sqrt{2\pi}\sigma} \int_{-\beta_k}^{+\infty} e^{-\frac{(z_k^{(i)})^2}{2\sigma^2}} dz_k^{(i)} \geq 1 - Pe_k, \quad (7)$$

as implied by (4). Thus the minimum scaling factor under the conventional ZF approach for the 4-QAM system is given by

$$\beta_k^- = -\sigma Q^{-1} \left(\sqrt{1 - Pe_k} \right), \quad (8)$$

where $Q^{-1}(\cdot)$ denotes the inverse of the standard Q -function [29]. When $M \geq 16$ (see Fig. 4 for 16-QAM), the standard constellation \mathcal{D}_S should be scaled so that

$$\frac{1}{\sqrt{2\pi}\sigma} \int_{-\beta_k}^{+\beta_k} e^{-\frac{(z_k^{(r)})^2}{2\sigma^2}} dz_k^{(r)} \times \frac{1}{\sqrt{2\pi}\sigma} \int_{-\beta_k}^{+\beta_k} e^{-\frac{(z_k^{(i)})^2}{2\sigma^2}} dz_k^{(i)} \geq 1 - Pe_k, \quad (9)$$

as implied by (4), considering the dominant scenario in which one of the inner most points is transmitted. Thus the minimum scaling factor under the conventional ZF approach is given by

$$\beta_k^- = \sigma Q^{-1} \left(0.5 - 0.5\sqrt{1 - Pe_k} \right). \quad (10)$$

In general, however, we do not have to zero-force \bar{y}_k , and in fact, it is sufficient to ensure \bar{y}_k falls into a region that favours correct symbol detection. This relaxation introduces room to optimize the choice of \mathbf{x} , leading to the following transmit power minimization problem:

$$\mathcal{P} : \begin{cases} \min_{\mathbf{x} \in \mathbb{C}^{N \times 1}} P(\mathbf{x}) = \mathbf{x}^\dagger \mathbf{x} \\ \text{subject to } \Pr((\bar{y}_k + z_k) \notin \mathcal{A}(d_k)) \leq Pe_k, \\ \text{for the transmitted data set } \{d_k \in \mathcal{D}_k, k = 1, \dots, K\}. \end{cases} \quad (11)$$

IV. SPARSE MC-CDMA WITH STANDARD M -QAM CONSTELLATIONS

In this section, we show how to translate the set of SEP targets in (11) into constraints on noiseless output components at MSs. In particular, we begin with the 4-QAM signaling case and then generalize to the 16-QAM signaling case. A similar approach can be applied to systems with higher-order QAM constellations.

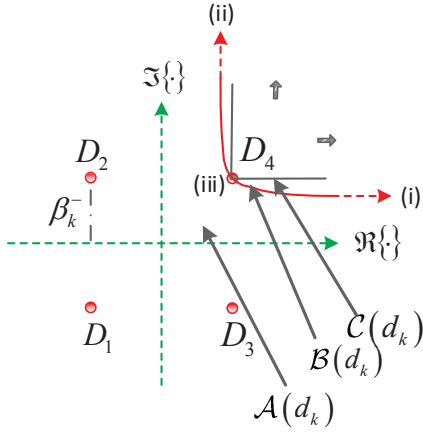


Fig. 3. 4-QAM: β_k^- is the minimum scaling factor under ZF; $\{D_1, D_2, D_3, D_4\}$ are constellation points; taking $d_k = D_4$ as an example, $\mathcal{A}(d_k)$ corresponds to the decision region, $\mathcal{B}(d_k)$ corresponds to the precise constraint region on noiseless output \bar{y}_m , while $\mathcal{C}(d_k)$ represents the constraint region with conservative approximation.

A. Translate SEP Targets to Constraints on Noiseless Received Signal Components

1) *4-QAM*: Assume that each d_k is drawn from the 4-QAM constellation set $\mathcal{D}_k = \{D_m : m = 1, \dots, 4\}$ as shown in Fig. 3, where the green dashed lines partition the complex plane into four symmetric decision regions each occupying an open quarter plane. Any received signals falling outside the correct region lead to detection error. Thus, the SEP requirement (4) becomes

$$\underbrace{\frac{1}{\sqrt{2\pi}\sigma} \int_{I_-^{(r)}(d_k)}^{I_+^{(r)}(d_k)} e^{-\frac{(z_k^{(r)})^2}{2\sigma^2}} dz_k^{(r)}}_{\mathcal{O}^{(r)}} \times \underbrace{\frac{1}{\sqrt{2\pi}\sigma} \int_{I_-^{(i)}(d_k)}^{I_+^{(i)}(d_k)} e^{-\frac{(z_k^{(i)})^2}{2\sigma^2}} dz_k^{(i)}}_{\mathcal{O}^{(i)}} \geq 1 - Pe_k, \quad (12)$$

where the tuple $(I_-^{(r)}(d_k), I_+^{(r)}(d_k), I_-^{(i)}(d_k), I_+^{(i)}(d_k))$ depends on the data transmitted and equals:

- (i) $(-\infty, -\bar{y}_k^{(r)}, -\infty, -\bar{y}_k^{(i)})$ for D_1 ;
- (ii) $(-\infty, -\bar{y}_k^{(r)}, -\bar{y}_k^{(i)}, +\infty)$ for D_2 ;
- (iii) $(-\bar{y}_k^{(r)}, +\infty, -\infty, -\bar{y}_k^{(i)})$ for D_3 ;
- (iv) $(-\bar{y}_k^{(r)}, +\infty, -\bar{y}_k^{(i)}, +\infty)$ for D_4 .

Given symbol $d_k \in \mathcal{D}_k$ and a target Pe_k , one can determine the precise constraint region $\mathcal{B}(d_k)$ on noiseless output \bar{y}_k at MS k from inequality (12). In particular, the boundary of the region is determined by equality in (12). For example, when $d_k = D_4$, three points on the boundary of $\mathcal{B}(d_k)$ are identified by considering combinations of $(\mathcal{O}^{(r)}, \mathcal{O}^{(i)})$:

- (i) $(1, 1 - Pe_k)$: $\bar{y}_k^{(r)} = +\infty, \bar{y}_k^{(i)} = -\sigma Q^{-1}(1 - Pe_k)$;
- (ii) $(1 - Pe_k, 1)$: $\bar{y}_k^{(r)} = -\sigma Q^{-1}(1 - Pe_k), \bar{y}_k^{(i)} = +\infty$;
- (iii) $(\sqrt{1 - Pe_k}, \sqrt{1 - Pe_k})$:
 $\bar{y}_k^{(r)} = \bar{y}_k^{(i)} = -\sigma Q^{-1}(\sqrt{1 - Pe_k})$.

In principle, the curved-shape boundary can be determined by traversing all possible combinations of $\mathcal{O}^{(r)}$ and $\mathcal{O}^{(i)}$. The constraint region $\mathcal{B}(d_k)$ then includes all the points on and within the boundary (see the red curve in Fig. 3). Note that point (iii) is exactly a scaled constellation point. Also note that this region generally implies non-linear constraints on input signal \mathbf{x} , which makes the optimization problem less tractable. Alternatively, we can find a polytype contained in $\mathcal{B}(d_k)$, i.e., we conservatively approximate the region using the area bounded by line segments between a finite number of points on or within boundary. A simple approximation for $d_k = D_4$ is given by: $\bar{y}_k^{(r)} \geq -\sigma Q^{-1}(\sqrt{1 - Pe_k})$ and $\bar{y}_k^{(i)} \geq -\sigma Q^{-1}(\sqrt{1 - Pe_k})$, which lead to linear constraints on input signal \mathbf{x} . This region with conservative approximation is denoted as $\mathcal{C}(d_k = D_4)$, see the open shadow area in Fig. 3.

By the same approach, relaxed constraint regions associated with the other constellation points can be derived and characterized as:

- (i) $\mathcal{C}(d_k = D_1) = \{(\bar{y}_k^{(r)}, \bar{y}_k^{(i)}) : \bar{y}_k^{(r)} \leq -I, \bar{y}_k^{(i)} \leq -I\}$;
- (ii) $\mathcal{C}(d_k = D_2) = \{(\bar{y}_k^{(r)}, \bar{y}_k^{(i)}) : \bar{y}_k^{(r)} \leq -I, -\bar{y}_k^{(i)} \leq -I\}$;
- (iii) $\mathcal{C}(d_k = D_3) = \{(\bar{y}_k^{(r)}, \bar{y}_k^{(i)}) : -\bar{y}_k^{(r)} \leq -I, \bar{y}_k^{(i)} \leq -I\}$,

with definition $I = -\sigma Q^{-1}(\sqrt{1 - Pe_k})$. Note that the exact areas of these regions only depend on the SEP targets.

2) *16-QAM*: We now turn to the case where d_k is drawn from the 16-QAM constellation set $\mathcal{D}_k = \{D_m : m = 1, \dots, 16\}$ as shown in Fig. 4. The constraint region on \bar{y}_k can be characterized based on procedures similar to those for the 4-QAM case. But calculation requires some care, since the decision regions $\mathcal{A}(d_k)$ for inner points and outer points are of different shapes (see the regions partitioned by the green dashed lines in Fig. 4) and the exact areas of these regions depend on the scaling factor β_k . For the sake of conciseness, the derivation of the constraint region $\mathcal{B}(d_k)$ is deferred to Appendix A. Taking symbols $\{D_{11}, D_{12}, D_{16}\}$ as examples, we plot the resulting regions $\mathcal{B}(d_k)$ in Fig. 4. Again, one can approximate these regions with a polytype $\mathcal{C}(d_k)$ for each, leading to linear constraints on input signal \mathbf{x} .

In general, the exact areas of $\mathcal{C}(d_k)$ depend on the scaling factor β_k . It is clear that, with minimum β_k^- as defined by (10), constraint region $\mathcal{C}(d_k)$ corresponds to a strict equality constraint on \bar{y}_k , i.e., $\bar{y}_k = D_m$, for the center points $m \in \{6, 8, 14, 16\}$; for the side points, e.g., D_{12} , $\mathcal{C}(d_k)$ shrinks to a line: $-\bar{y}_k^{(r)} \leq -2\beta_k^- - I$ and $\bar{y}_k^{(i)} = \beta_k^-$, which implies doing a zero-forcing for the imaginary part while having a relaxed constraint on the real part; and for the corner points, e.g., D_{11} , $\mathcal{C}(d_k)$ becomes: $-\bar{y}_k^{(r)} \leq -2\beta_k^- - I, -\bar{y}_k^{(i)} \leq -2\beta_k^- - I$, where we recall that $I = -\sigma Q^{-1}(\sqrt{1 - Pe_k})$.

To meet a given SEP target Pe_k for MS k , one could certainly adopt a scaling factor larger than the minimum β_k^- for the transmission. But such a choice may affect the power efficiency of the system. Fig. 5 plots two instances of constraint regions $\mathcal{B}(d_k)$ for $Pe_k = 10^{-3}$ when the constellation is scaled up with $\beta_k = 1.05\beta_k^-$ (blue curves) and $\beta_k = 1.20\beta_k^-$ (red curves). It can be seen that with a larger β_k above β_k^- , when a center constellation point is transmitted, the constraint region is relaxed from a single point to a circle-

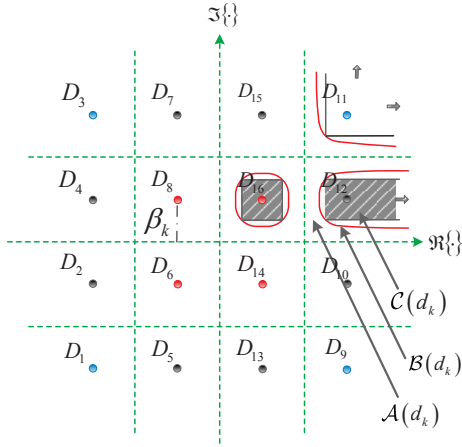


Fig. 4. 16-QAM: β_k is a scaling factor; $\{D_1, \dots, D_{16}\}$ are constellation points; taking $d_k = D_{12}$ as an example, $\mathcal{A}(d_k)$ corresponds to the decision region, $\mathcal{B}(d_k)$ corresponds to the precise constraint region on noiseless output \bar{y}_k , while $\mathcal{C}(d_k)$ represents the constraint region with conservative approximation.

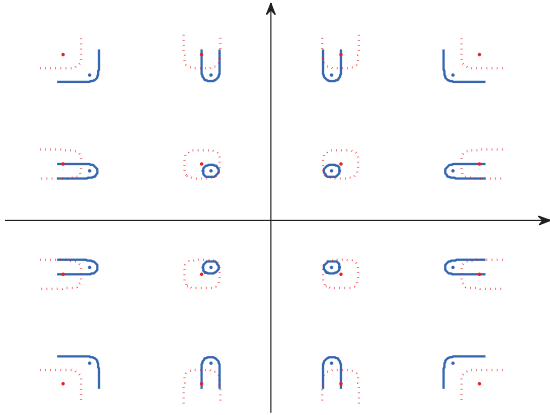


Fig. 5. Constraint regions $\mathcal{B}(d_k)$ for the 16-QAM constellation under different scaling factors with $Pe_k = 10^{-3}$ fixed: (i) blue curves: $\beta_k = 1.05\beta_k^-$; (ii) red curves: $\beta_k = 1.20\beta_k^-$.

type region centered on the symbol. Potential benefits can be accrued from the resulting enlarged feasible region. However, when the corner points are transmitted, the corresponding constraint regions always shrink as the constellation is scaled up. In this case, power efficiency loss may be induced because of the reduced feasible optimization space. When one of the side points is transmitted, it is unclear how the performance reacts, as the constraint region with larger β_k partially overlaps with that for a smaller β_k . Nevertheless our experiments have indicated that when the transmitted data symbols are randomly and uniformly generated, scaling up the constellation with $\beta_k > \beta_k^-$ brings little further power saving. Hence we will use the minimum scaling, β_k^- , for the standard 16-QAM in what follows.

B. Problem Reformulation with Conservative Approximation

With the conservative approximation, the SEP constraints in (11) can be translated into a set of linear inequality/equality

constraints on vector \mathbf{x} . We now present the resulting optimization problem.

For ease of exposition, we stack the real and the imaginary parts of each x_n into a real vector $\tilde{\mathbf{x}}$, i.e., $\tilde{\mathbf{x}} = [\Re\{x_1\}, \Im\{x_1\}, \dots, \Re\{x_N\}, \Im\{x_N\}]^T \in \mathbb{R}^{2N \times 1}$. For the 4-QAM signaling, the real and imaginary parts of coded signal x_n are associated with different inequalities. For the 16-QAM signaling, with $\beta_k = \beta_k^-$, the real and the imaginary parts of x_n are associated with either one equality or an inequality depending on the data symbols; when $\beta_k > \beta_k^-$, at most two inequalities are introduced for either the real or the imaginary part. Therefore, the optimization problem of (11) can be generally reformulated as follows:

$$\mathcal{P}_1 : \begin{cases} \min_{\tilde{\mathbf{x}} \in \mathbb{R}^{2N \times 1}} P(\tilde{\mathbf{x}}) = \tilde{\mathbf{x}}^T \tilde{\mathbf{x}} \\ \text{s. t. } \mathbf{A}\tilde{\mathbf{x}} - \mathbf{c} \preceq \mathbf{0}, \\ \mathbf{B}\tilde{\mathbf{x}} - \mathbf{e} = \mathbf{0}, \end{cases} \quad (13)$$

where for the 4-QAM system, we have $\mathbf{A} = \{a_{i,j}\} \in \mathbb{R}^{2K \times 2N}$, $\mathbf{c} = \{c_i\} \in \mathbb{R}^{2K \times 1}$, $\mathbf{B} = \mathbf{0}$ and $\mathbf{e} = \mathbf{0}$, while for the 16-QAM system, we have $\mathbf{A} = \{a_{i,j}\} \in \mathbb{R}^{4K \times 2N}$, $\mathbf{c} = \{c_i\} \in \mathbb{R}^{4K \times 1}$, $\mathbf{B} = \{b_{i,j}\} \in \mathbb{R}^{2K \times 2N}$ and $\mathbf{e} = \{e_i\} \in \mathbb{R}^{2K \times 1}$. The precise definitions of the matrices and the vectors involved depend on constraint regions $\mathcal{C}(d_k)$, $\forall k$, as defined in Section IV-A1 and Section IV-A2 for the 4-QAM and 16-QAM signalling, respectively. Whether or not the inequality/equality constraints are active will depend on the transmitted data associated with the constellation. It is possible that some of the constraints may not be active, in which case the corresponding row entries of \mathbf{A}/\mathbf{B} and \mathbf{c}/\mathbf{e} are padded with zeros. In addition, matrices \mathbf{A}/\mathbf{B} enjoy the same sparsity as the system matrix \mathbf{H} . In particular, when indexing matrices \mathbf{A}/\mathbf{B} , the index pair (i, j) corresponds to a particular MS-subcarrier pair, and thus $a_{i,j}/b_{i,j} = 0$ whenever the particular subcarrier is not allocated to the MS.

As an illustrative example, we consider a 16-QAM system with $N = 3$ subcarriers and $K = 2$ MSs. Each MS is allocated $L = 2$ subcarriers: The first MS is allocated subcarriers 1 and 2, while the second MS is allocated subcarriers 2 and 3. The effective channel matrix \mathbf{H} is assumed to be

$$\mathbf{H} = \begin{bmatrix} 1+j & -1+j & 0 \\ 0 & 1+j & 1-j \end{bmatrix}. \quad (14)$$

The data symbol intended for MS 1 and 2 is $d_1 = D_{16}$ and $d_2 = D_{11}$, respectively. Both MSs require the same SEP target Pe and thus employ the same scaling factor $\beta = \beta_1 = \beta_2$. The noiseless received components at MSs are calculated as

$$\begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \end{bmatrix} = \mathbf{H}\mathbf{x} = \begin{bmatrix} (1+j)x_1 + (-1+j)x_2 \\ (1+j)x_2 + (1-j)x_3 \end{bmatrix}. \quad (15)$$

Therefore, according to the constraint regions constructed (see Section IV-A2 and also Appendix A), we have:

$$\beta - \delta_0 \leq \bar{y}_1^{(r)} \leq \beta + \delta_0, \quad \beta - \delta_0 \leq \bar{y}_1^{(i)} \leq \beta + \delta_0, \quad (16)$$

$$-\bar{y}_2^{(r)} \leq -\left(2\beta - \sigma Q^{-1} \left(\sqrt{1 - Pe}\right)\right),$$

$$-\bar{y}_2^{(i)} \leq -\left(2\beta - \sigma Q^{-1} \left(\sqrt{1 - Pe}\right)\right), \quad (17)$$

which, by simple algebra, can be further translated into a set of general linear constraints as in (13) with $\mathbf{B} = \mathbf{0}$, $\mathbf{e} = \mathbf{0}$,

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & -1 & -1 & 0 & 0 \\ -1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & -1 & 0 & 0 \\ -1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\text{and } \mathbf{c} = \begin{bmatrix} \beta + \delta_0 \\ -(\beta - \delta_0) \\ \beta + \delta_0 \\ -(\beta - \delta_0) \\ -(2\beta - \sigma Q^{-1}(\sqrt{1 - Pe})) \\ 0 \\ -(2\beta - \sigma Q^{-1}(\sqrt{1 - Pe})) \\ 0 \end{bmatrix}. \quad (18)$$

As we can observe, matrix \mathbf{A} inherits sparsity from the effective channel matrix \mathbf{H} ; transmission of $d_1 = D_{16}$ and $d_2 = D_{11}$ invokes four and two inequality constraints, respectively, and there is no strict equality constraint in this example.

V. PRECODING ALGORITHM DESIGN WITH PARALLEL COMPUTATION UNITS

Problem \mathcal{P}_1 is strictly convex and can be solved via a number of standard algorithms, such as the interior-point algorithm [30]. Most of these algorithms are designed for centralized implementation and could be efficient enough for a small-scale problem. However, as the problem dimension increases, the computational complexity may be prohibitive. More importantly, the sparsity inherent to the problem may not be well exploited in standard solvers. Given these observations, we are interested in developing a precoding algorithm that leverages the sparsity to reduce complexity and is suitable for solving large-dimension problems using parallel computation units. These units may correspond to parallel processor cores (threads) at the BS computer [31] or parallel processors at the cloud to which the BS is connected [32].

A. Algorithm Design

We now detail the precoding algorithm with a focus on the general problem as if all inequalities and equalities in (13) were activated. The key technique is the dual decomposition approach, see, e.g., [33]. Recalling the graph representation \mathcal{G} introduced (see Fig. 2), we can map all PSNs and OSNs to parallel computation units.

We start with forming the Lagrangian function:

$$\mathcal{L}(\tilde{\mathbf{x}}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \tilde{\mathbf{x}}^T \tilde{\mathbf{x}} + \boldsymbol{\lambda}^T (\mathbf{A}\tilde{\mathbf{x}} - \mathbf{c}) + \boldsymbol{\nu}^T (\mathbf{B}\tilde{\mathbf{x}} - \mathbf{e}), \quad (19)$$

where $\boldsymbol{\lambda} \in \mathbb{R}^{4K \times 1} \succeq 0$ and $\boldsymbol{\nu} \in \mathbb{R}^{2K \times 1}$ are Lagrangian multipliers (dual variables), among which each pair of primal variables $\tilde{\mathbf{x}}_{[2n-1:2n]}$ is associated with PSN computation unit n ($n = 1, \dots, N$), while each tuple of dual variables $\{\boldsymbol{\lambda}_{[4k-3:4k]}, \boldsymbol{\nu}_{[2k-1:2k]}\}$ is associated with OSN computation unit k ($k = 1, \dots, K$). The dual problem is then defined as

$$\max_{\boldsymbol{\lambda}, \boldsymbol{\nu}} g(\boldsymbol{\lambda}, \boldsymbol{\nu}), \quad \text{subject to } \boldsymbol{\lambda} \succeq 0, \quad (20)$$

with $g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \min_{\tilde{\mathbf{x}}} \mathcal{L}(\tilde{\mathbf{x}}, \boldsymbol{\lambda}, \boldsymbol{\nu})$ being the dual objective function.

Then one can solve the original problem by finding the optimal dual variables in an iterative manner. Specifically, at the t th iteration, for fixed dual variables $\boldsymbol{\lambda}^{(t)}$ and $\boldsymbol{\nu}^{(t)}$, to attain the minimization of Lagrangian, one sets the first-order derivative of the Lagrangian to zero, which leads to

$$\tilde{\mathbf{x}}^{*(t)} = -\frac{1}{2} (\mathbf{A}^T \boldsymbol{\lambda}^{(t)} + \mathbf{B}^T \boldsymbol{\nu}^{(t)}), \quad (21)$$

or more explicitly,

$$\tilde{x}_l^{*(t)} = -\frac{1}{2} \left(\sum_{i \in \mathcal{I}(\tilde{x}_l)} a_{i,l} \lambda_i^{(t)} + \sum_{i \in \mathcal{I}'(\tilde{x}_l)} b_{i,l} \nu_i^{(t)} \right), \quad l = 1, \dots, 2N, \quad (22)$$

where $\mathcal{I}(\tilde{x}_l)$ and $\mathcal{I}'(\tilde{x}_l)$ denote the collection of indices of dual variables λ and ν that have interactions with \tilde{x}_l , respectively, according to the graph \mathcal{G} . Note that PSN unit n is in charge of computing the pair $\tilde{\mathbf{x}}_{[2n-1:2n]}^{*(t)}$, $n = 1, \dots, N$. The corresponding dual function is given by:

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \tilde{\mathbf{x}}^{*(t)T} \tilde{\mathbf{x}}^{*(t)} + \boldsymbol{\lambda}^T (\mathbf{A}\tilde{\mathbf{x}}^{*(t)} - \mathbf{c}) + \boldsymbol{\nu}^T (\mathbf{B}\tilde{\mathbf{x}}^{*(t)} - \mathbf{e}). \quad (23)$$

The dual variables are then updated by the OSN units in a parallel manner according to

$$\begin{aligned} \lambda_i^{(t+1)} &= \left[\lambda_i^{(t)} + \frac{t-1}{t+2} (\lambda_i^{(t)} - \lambda_i^{(t-1)}) \right. \\ &\quad \left. + \frac{1}{2\kappa} \left(\sum_{l \in \mathcal{I}(\lambda_i)} a_{i,l} \hat{x}_l^{(t)} - c_i \right) \right]^+, \quad i = 1, \dots, 4K, \\ \nu_j^{(t+1)} &= \nu_j^{(t)} + \frac{t-1}{t+2} (\nu_j^{(t)} - \nu_j^{(t-1)}) \\ &\quad + \frac{1}{2\kappa} \left(\sum_{l \in \mathcal{I}(\nu_j)} b_{j,l} \hat{x}_l^{(t)} - e_j \right), \quad j = 1, \dots, 2K, \end{aligned} \quad (24)$$

where $\mathcal{I}(\lambda_i)$ and $\mathcal{I}(\nu_j)$ denote the collection of indices of primal variables that have interactions with λ_i and ν_j , respectively; notation $[\cdot]^+$ denotes the projection onto the nonnegative orthant, $\kappa = (\|\bar{\mathbf{A}}\bar{\mathbf{A}}^T\|_1 \|\bar{\mathbf{A}}\bar{\mathbf{A}}^T\|_\infty)^{1/2}$ with $\bar{\mathbf{A}} = [\mathbf{A}^T, \mathbf{B}^T]^T$ and

$$\hat{x}_l^{(t)} = \tilde{x}_l^{*(t)} + \frac{t-1}{t+2} (\tilde{x}_l^{*(t)} - \tilde{x}_l^{*(t-1)}), \quad l = 1, \dots, 2N.$$

This dual-variable updating rule offers faster convergence speed than the conventional gradient updating [33] as shown in [34]. The algorithm described is summarized in Table I.

B. Complexity Analysis

To quantify the complexity of the algorithm, we distinguish the communication overhead and computational complexity.

In the algorithm, to update its primal variables $\tilde{\mathbf{x}}_{[2n-1:2n]}$ via (22), PSN unit n only needs to gather dual variables from its neighboring OSNs; therefore, the number of messages passed to PSN unit n depends on the number of active dual variables and is at most $4|\mathcal{I}(x_n)|$ for the 16-QAM, and $2|\mathcal{I}(x_n)|$ for the 4-QAM. To update dual variables $\{\boldsymbol{\lambda}_{[4k-3:4k]}, \boldsymbol{\nu}_{[2k-1:2k]}\}$, OSN unit k only needs to collect

TABLE I
PRECODING ALGORITHM WITH PARALLEL COMPUTATION UNITS

1) Initialize: dual variables $\lambda^{(0)} > 0, \nu^{(0)}$;
2) Repeat for T iterations until convergence criterion is met:
2.1) For $n = 1, \dots, N$:
PSN unit n computes its primal variables $\tilde{\mathbf{x}}_{[2n-1:2n]}$ using (22), and broadcasts the updated values to its neighboring OSN units;
2.2) For $k = 1, \dots, K$:
OSN unit k computes its dual variables $\{\lambda_{[4k-3:4k]}, \nu_{[2k-1:2k]}\}$ using (24) and broadcasts the updated values to its neighboring PSN units.

primal variables from its neighboring PSNs; therefore, the number of messages passed to OSN unit k is at most $2|\mathcal{I}(y_k)|$ for both the 4/16-QAM. Thus, with T iterations, the total number of message-passings across computation units is $\mathcal{O}(4KLT)$ and $\mathcal{O}(6KLT)$ for the 4-QAM and 16-QAM system, respectively, where we have used the fact that $\sum_{n=1}^N |\mathcal{I}(x_n)| = \sum_{k=1}^K |\mathcal{I}(y_k)| = KL$ with L being the number of non-zeroes in each signature.

Table II summarizes the computational complexity for the algorithm proposed. To update each of its primal variables via (22), PSN unit n needs at most $(4|\mathcal{I}(x_n)| - 1)$ additions and $4|\mathcal{I}(x_n)|$ multiplications for the 16-QAM, and $(2|\mathcal{I}(x_n)| - 1)$ additions and $2|\mathcal{I}(x_n)|$ multiplications for the 4-QAM. On the other hand, to update each of its dual variables via (24), OSN unit k needs at most $(6|\mathcal{I}(y_k)| + 5)$ additions and $(4|\mathcal{I}(y_k)| + 2)$ multiplications for both the 4/16-QAM. Overall, the algorithm involves $\mathcal{O}(16KLT)$ additions and multiplications for the 4-QAM system, and involves $\mathcal{O}(20KLT)$ additions and multiplications for the 16-QAM system, where T is the number of iterations. It is clear that the more sparse the signatures are, the less overall communication overhead and computational complexity are required to generate the precoded symbols in proposed algorithm.

For comparison, we note that the conventional ZF precoding of (5) has a computational complexity of $\mathcal{O}(\frac{8}{3}K^3 + 4NK^2)$ to compute precoding matrix $\mathbf{W} = \mathbf{H}^\dagger(\mathbf{H}\mathbf{H}^\dagger)^{-1}$ and additional complexity of $\mathcal{O}(4KN)$ to generate each precoded symbol vector via $\mathbf{x} = \mathbf{W}\mathbf{d}$. Consider a transmission frame that consists of T_s 4-QAM symbols intended for each MS and assume channel \mathbf{H} remains unchanged during the frame. The ratio of the complexity of the proposed scheme to that of the ZF approach is thus quantified by $\rho = \mathcal{O}(16KLT_s)/\mathcal{O}(\frac{8}{3}K^3 + 4NK^2 + 4KNT_s)$. It is clear that the smaller L , the smaller ρ will be. In particular, $\rho \approx 4LT/K$ for a fully loaded system with $N = K$ and sufficiently large T_s . As an example with $K = 32$, $L = 4$ and $T = 100$ iterations, ratio $\rho \approx 50$, which indicates the proposed scheme has approximately 50 times the complexity of ZF precoding. Despite the increase in complexity, the proposed scheme is able to provide enormous transmit power reduction over ZF precoding and is more robust against imperfect channel state estimation, as will be shown later in Section VII.

VI. SPARSE MC-CDMA WITH REPLICA CONSTELLATIONS

In the previous sections, we have mainly focused on the system with standard QAM constellations for which convex constraints on the precoded vector are constructed according to the SEP targets. In this section, we assume the system adopts replica constellations, where each \mathcal{D}_k is the periodic extension of a regular QAM constellation along the real and imaginary axes. We propose an optimized Tomlinson-Harashima Precoding (THP) under the SEP constraints by applying a similar approach we have used for the system with standard QAM constellations.

A. THP-Basics

We first briefly review some basic concepts related to THP, see, e.g., [35].

In general, the replica constellation point $d \in \mathcal{D}_k$ can be represented as:

$$d = D_l + 2\beta_k\sqrt{M}(a_R + ja_I), \quad (25)$$

where D_l corresponds to a regular point in the scaled M -QAM constellation under scaling β_k ($l = 1, \dots, M$), and $\{a_R, a_I\}$ corresponds to an arbitrary integer pair, see Fig. 6 for a visual illustration when $M = 4$. It is noted that decision regions associated with all replica points are identical closed squares with side length $2\beta_k$.

The THP is normally done in a successive manner in which interference created by previous users' transmissions is pre-cancelled to facilitate the transmission for the current user at each stage. The encoding is accommodated by the replica constellation and modulo-operation at the transmitter. Specifically, let the channel matrix be represented as $\mathbf{H}^\dagger = \mathbf{F}\mathbf{R}$ as a result of QR factorization, where \mathbf{F} is a unitary matrix and \mathbf{R} is an upper triangular matrix. Then $\mathbf{B} = \mathbf{H}\mathbf{F} = \mathbf{R}^\dagger$ is a lower triangular matrix. The successive precoding operates as

$$\bar{x}_k = \frac{1}{\mathbf{B}_{k,k}} \left[d_k - \sum_{l=1}^{k-1} \mathbf{B}_{k,l} \bar{x}_l \right]_{p_k}, \quad (26)$$

where $d_k \in \mathcal{D}_k$ is the replica point carrying information for MS k and $[u]_{p_k}$ is the modulo operation operated on complex number u with respect to basis p_k and is defined as:

$$[u]_{p_k} = u - \left\lfloor \frac{\Re\{u\} + p_k/2}{p_k} \right\rfloor p_k - j \left\lfloor \frac{\Im\{u\} + p_k/2}{p_k} \right\rfloor p_k, \quad (27)$$

with $p_k = \sqrt{M}\beta_k$. The transmit signal is then formed by multiplying \mathbf{F} with $\bar{\mathbf{x}}$, i.e., $\mathbf{x} = \mathbf{F}\bar{\mathbf{x}}$. In this way, at the receiver side, no MS experiences inter-user interference because of the pre-cancellation operations done at the BS. It is remarked that since THP is performed in a successive manner, different user orderings may lead to different performance. To find the optimal ordering, one needs to do an exhaustive search over all possible combinations, which is generally infeasible as K goes large. In this work we simply adopt the suboptimal V-BLAST (VB) ordering [36].

TABLE II
 COMPUTATIONAL COMPLEXITY OF THE PROPOSED ALGORITHM

Schemes	Operations	"+" / iteration	"×" / iteration	complexity (T iterations)
4-QAM	compute (22) :	$4KL - 2N$	$4KL$	$\mathcal{O}(16KLT)$
	compute (24) :	$12KL + 10K$	$8KL + 4K$	
16-QAM	compute (22) :	$8KL - 2N$	$8KL$	$\mathcal{O}(20KLT)$
	compute (24) :	$12KL + 10K$	$8KL + 4K$	

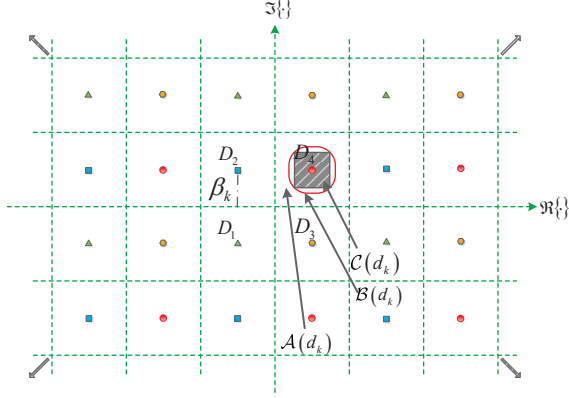


Fig. 6. Replica constellation built on 4-QAM: β_k is a scaling factor; $\{D_1, D_2, D_3, D_4\}$ correspond to regular signal points before periodic extension; taking $d_k = D_4$ as an example, $A(d_k)$ corresponds to the decision region, $B(d_k)$ corresponds to the precise constraint region on noiseless output \bar{y}_m , while $C(d_k)$ represents the constraint region with conservative approximation.

B. Optimized THP

Under the SEP constraints, we can formulate a precoding optimization problem similar to \mathcal{P}_1 . The idea is that instead of choosing a minimum scaling factor β_k^- for the constellation and performing a zero-forcing THP, one can relax the constraints on the noiseless output components and introduces room for optimizing the input signals as we scale up the constellation.

In particular, with the replica constellation scaled up, the resulting constraint regions become boxes centered at each replica point as constructed and approximated similarly for the inner points of the 16-QAM constellation, see Fig. 6. The optimization problem is then formulated as:

$$\mathcal{P}_2 : \begin{cases} \min_{\tilde{\mathbf{x}} \in \mathbb{R}^{2N \times 1}} P(\tilde{\mathbf{x}}) = \tilde{\mathbf{x}}^T \tilde{\mathbf{x}} \\ \text{s. t. } \mathbf{A}\tilde{\mathbf{x}} - \mathbf{c} \leq \mathbf{0}, \end{cases} \quad (28)$$

where matrices $\mathbf{A} = \{a_{i,j}\} \in \mathbb{R}^{4K \times 2N}$ and $\mathbf{c} = \{c_i\} \in \mathbb{R}^{4K \times 1}$ are formed according to the constraints

$$\mathcal{C}(d_k) = \left\{ \left(\bar{y}_k^{(r)}, \bar{y}_k^{(i)} \right) : \begin{aligned} d_k^{(r)} - \delta_0 &\leq \bar{y}_k^{(r)} \leq d_k^{(r)} + \delta_0 \\ d_k^{(i)} - \delta_0 &\leq \bar{y}_k^{(i)} \leq d_k^{(i)} + \delta_0 \end{aligned} \right\}, \quad k = 1, \dots, K, \quad (29)$$

where parameter δ_0 determines the size of the constraint box and is chosen to satisfy:

$$Q\left(\frac{\delta_0 - \beta_k}{\sigma}\right) - Q\left(\frac{\delta_0 + \beta_k}{\sigma}\right) = \sqrt{1 - Pe_k}, \quad (30)$$

and the set of information-carrying replica points $\{d_k, k = 1, \dots, K\}$ is determined from the ZF-THP encoding procedure. Therefore, the precoding algorithm proposed in Section V can be applied here to calculate the optimized THP precoded vector.

VII. SIMULATION RESULTS

We now present numerical results to demonstrate the effectiveness of the precoding schemes proposed for the sparse MC-CDMA system.

A. Simulation Setup

In the simulation, the noise variance N_0 is set to unity. The total number of subcarriers is fixed with $N = 32$ and the number of MSs $K \leq N$ is allowed to vary. Different MSs experience different frequency-selective fading channels. Specifically, the channel frequency response between the BS and MS k is generated according to

$$\tilde{h}_{k,n} = \sum_{q=0}^{\tilde{Q}-1} g_{k,q} e^{-j \frac{2\pi q n}{N}}, \quad n = 1, \dots, N, \quad (31)$$

where $\mathbf{g}_k = [g_{k,0}, \dots, g_{k,\tilde{Q}-1}]^T$ represents the discrete-time channel response consisting of \tilde{Q} taps; components $\{g_{k,q}\}$ are modeled as independent zero-mean Gaussian random variables, whose individual variance equals $\{\bar{\lambda} e^{-\frac{q}{\tilde{Q}}}\}$ with normalization factor $\bar{\lambda}$ chosen such that $E[\|\mathbf{g}_k\|^2] = 1$. In the simulation, $\tilde{Q} = 8$ is adopted for the channel generation. It is assumed that there is no inter-symbol-interference and inter-carrier-interference in the system.

Unless stated otherwise, for any fixed system configuration, we simulate 1000 transmission slots, under each of which random data and random channel are independently generated for each MS. In addition, we produce 10 random regular signature matrix realizations as defined in Section II-C and thus every 100 transmission slots share the same signature matrix. The system transmitted power consumption presented shortly is averaged over all transmission slots. For the precoding algorithm proposed, the calculation terminates at iteration t if the normalized improvement of dual objective $|g^{(t)} - g^*|/g^* \leq \delta = 10^{-4}$, where $g^* \triangleq \max_{i \in \{1, \dots, t-1\}} g^{(i)}$.

B. System with Standard Constellation

Fix SEP target $Pe_k = 10^{-3}$ for all MSs and vary $K \in [24 : 32]$. In the sparse MC-CDMA, different levels of sparsity, e.g., $L = 4$ and $L = 8$, are considered. The case with $L = 32$, referred to as the dense MC-CDMA, is also considered for the

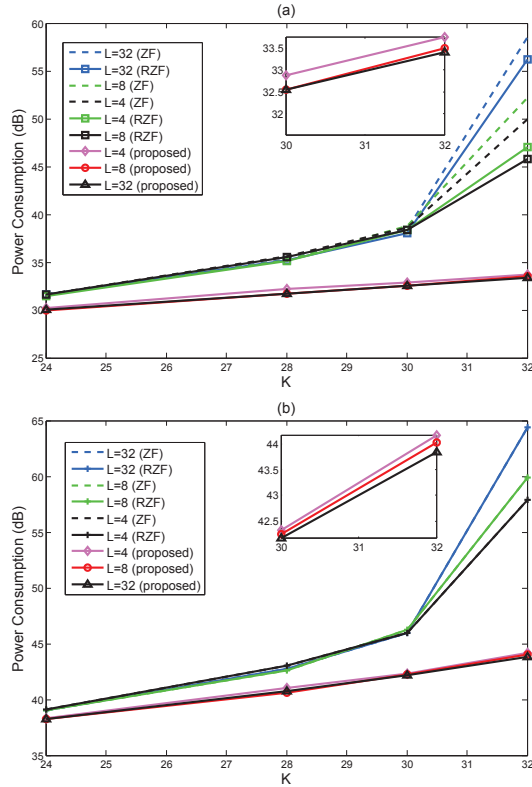


Fig. 7. Transmit power consumption at BS under different loads and different levels of sparsity ($N = 32$, $Pe_k = 10^{-3}$): (a) standard 4-QAM; (b) standard 16-QAM.

purpose of comparison. In addition, we have also compared with two conventional precoding schemes including ZF and the optimized RZF in form of:

$$\mathbf{x}_{ZF} = \mathbf{H}^\dagger (\mathbf{H}\mathbf{H}^\dagger)^{-1} \mathbf{d}, \quad (32)$$

$$\mathbf{x}_{RZF} = k_1 \mathbf{H}^\dagger (\mathbf{H}\mathbf{H}^\dagger + k_2 \mathbf{I}_K)^{-1} \mathbf{d}, \quad (33)$$

where $\mathbf{d} = [d_1, \dots, d_K]^T$ with d_k denotes the transmitted data symbol for MS k and is drawn from a scaled version of standard constellation by β_k^- , \mathbf{I}_K is a $K \times K$ identity matrix, and k_1, k_2 are two non-negative parameters to be optimized subject to the SEP constraints in (13). Note that the optimized RZF encompasses the conventional regularized ZF precoder [12] with $k_1 = 1$ and also the minimum-mean-square-error (MMSE) precoder [20] with $k_1 = 1$ and $k_2 = K\sigma^2/P$.

Fig. 7-(a) and Fig. 7-(b) plot the transmit power consumption at the BS versus K under different setups and precoding schemes for systems with 4-QAM and 16-QAM, respectively. Two important observations are made as follows.

First, power reduction from the proposed precoding over ZF and RZF is clearly evident for all load and sparsity combinations considered. In particular, for any fixed L , the reduction increases as K grows. For instance, when $K = N = 32$ and $L = 8$, for the 4-QAM, we have 18.8 dB and 13.6 dB reduction compared with ZF and RZF, respectively, while for the 16-QAM, we have 16 dB reduction compared with both schemes, noting that the optimized RZF solutions are degraded

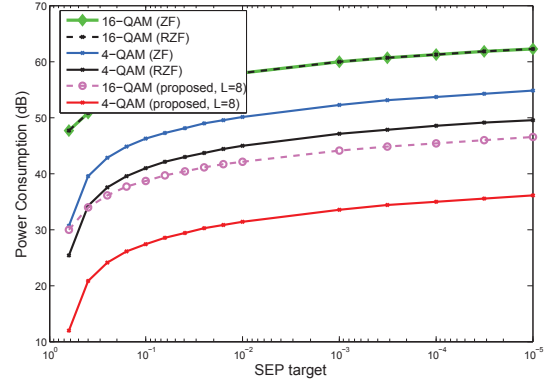


Fig. 8. Power consumption versus different SEP targets ($K = N = 32$, $L = 8$).

and coincide with the ZF solutions in this case. As $K \rightarrow N$, the effective channel matrix \mathbf{H} is increasingly likely to be poorly conditioned. Hence, the inefficiency of conventional schemes (in particular ZF) becomes pronounced. However, the precoding proposed is not sensitive to the conditional number of \mathbf{H} and always attains the best performance.

Second, for the sparse MC-CDMA system, there is a trend that a denser signature (a larger value of L) leads to a smaller power consumption needed. For instance, the system with $L = 4$ consumes slightly more power than the system with $L = 8$ to achieve the same SEP target under both the 4-QAM and 16-QAM systems. However, to attain comparable power efficiency to the dense MC-CDMA, the signatures can still be relatively sparse ($L = 8$ in our examples), yielding considerable reduction in precoding complexity. This observation also indicates that the sparse MC-CDMA system with proper choice of L would attain almost the same link throughput as that of the dense MC-CDMA system under the same transmit power budget.

Fig. 8 plots the power consumption versus different SEPs with $K = N = 32$ and $L = 8$, which further confirms the superiority of the proposed scheme as compared to baselines ZF and RZF under different SEP targets for both 4-QAM and 16-QAM systems.

C. System with Replica Constellation

We now consider the system with replica constellation, where system parameters $N = 32$ and $L = 8$. All MSs request the same minimum SEP target $Pe = 10^{-3}$. Under this SEP requirement, a uniform scaling factor across all MSs is chosen such that the power consumption is minimized for the proposed optimized THP. To perform the precoding optimization, we use an algorithm similar to that for systems with standard constellations. Therefore, signature sparsity is leveraged to reduce precoding complexity, as it was before.

Fig. 9-(a) and Fig. 9-(b) plot the transmit power consumption versus K under both the ZF- and optimized THP schemes for the system with 4-QAM and 16-QAM replica points, respectively. The performance of the proposed scheme under standard constellations is also included here for the purpose of comparison. It is seen that the optimized THP is

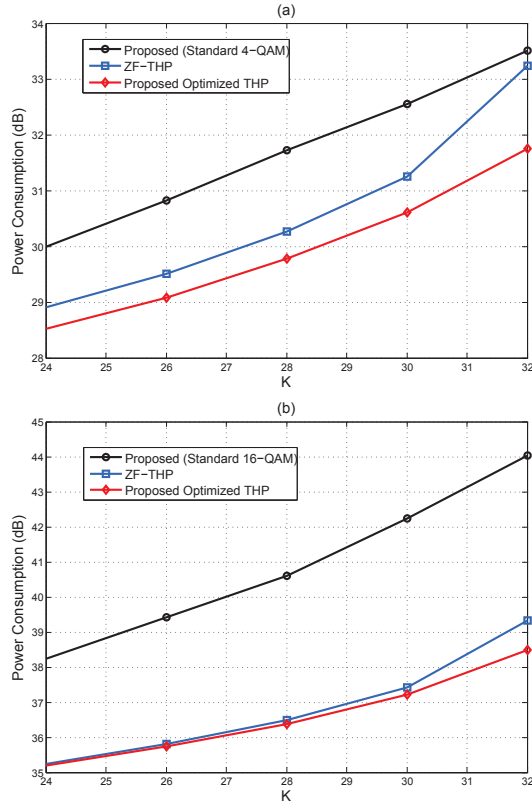


Fig. 9. Power consumption for ZF/optimized THP with replica points under different loads: (a) extended from standard 4-QAM; (b) extended from standard 16-QAM ($N = 32, L = 8, Pe_k = 10^{-3}$).

able to provide significant power reduction over the proposed scheme under standard constellations. This further reduction, albeit appealing, does not come for free and has to be paid with more sophisticated encoding and decoding operations in THP schemes as described in Section VI. It is also observed that the optimized THP generally outperforms ZF-THP in power efficiency and the exact gain depends on the system load. In particular, the former provides roughly 1.5 dB power reduction over the latter for 4-QAM replica and roughly 0.85 dB reduction for 16-QAM replica in a full-load system. The ZF-THP is already very power-efficient, yet the proposed THP is seen here to provide further reduction in transmit power.

D. Bit Error Rate (BER) Results and Impact of Imperfect Channel Estimation

So far, we have demonstrated the power efficiency of the proposed precoding under different uncoded SEP targets. We now evaluate the impact of the proposed scheme on another practically important performance metric in terms of uncoded bit error rate (BER). The BER is calculated and averaged over 10^6 realizations of transmissions. Fig. 10 (a) and (b) depict the average BER (at a typical MS) as a function of power consumption for a system with standard/replica 4-QAM and 16-QAM, respectively. Consistent with the previous observations, the proposed optimized precoding significantly outperforms ZF precoding in terms of power efficiency to attain the same BER target under standard QAM constellations. The optimized

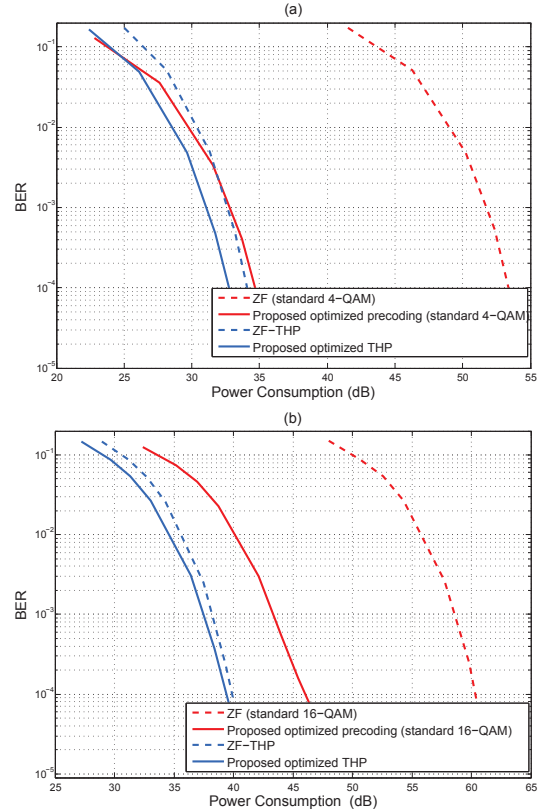


Fig. 10. Uncoded BER versus power consumption ($K = N = 32, L = 8$): (a) standard and replica 4-QAM; (b) standard and replica 16-QAM.

THP is more power-efficient than the ZF-THP, and both of them generally outperform the optimized precoding under standard constellations but at the cost of increase complexity as explained before.

Next, we evaluate the impact of channel estimation error (i.e., channel uncertainty) on the performance of the schemes considered. Let $\hat{\mathbf{H}}$ denote the estimated sparse channel matrix. Each nonzero entry $\hat{h}_{k,n}$ of $\hat{\mathbf{H}}$ is a noisy version of the perfect $h_{k,n}$ of \mathbf{H} . To model the uncertainty, we assume $\hat{h}_{k,n}$ is generated according to: $\hat{h}_{k,n} = h_{k,n} + z_{k,n}$, where $z_{k,n} \sim \mathcal{CN}(0, \sigma_e^2)$ represents the complex Gaussian estimation error with variance σ_e^2 and $\{z_{k,n}, \forall k, \forall n\}$ are independently and identically distributed. The average normalized channel uncertainty is then defined as $\tau = \mathbb{E}[10 \log_{10}(\|\hat{\mathbf{H}} - \mathbf{H}\|_2^2 / \|\mathbf{H}\|_2^2)]$ in dB. For all schemes evaluated, the SEP target Pe_k is set to 10^{-2} so that the corresponding BER is on the order of 10^{-3} , if perfect channel state information is available. Fig. 11 (a) and (b) depict the real BER (at a typical MS) versus different levels of channel uncertainty for a sparse MC-CDMA system with standard/replica 4-QAM and 16-QAM signaling, respectively. It can be seen that as the channel uncertainty increases, the real BER of both the ZF approach and the proposed precoding scheme degrade. However, the proposed scheme always outperforms its ZF counterpart and exhibits much better robustness against imperfect channel estimation, particularly for a system with standard constellations. The proposed scheme is thus not only more power-efficient but also more robust against channel uncertainty.

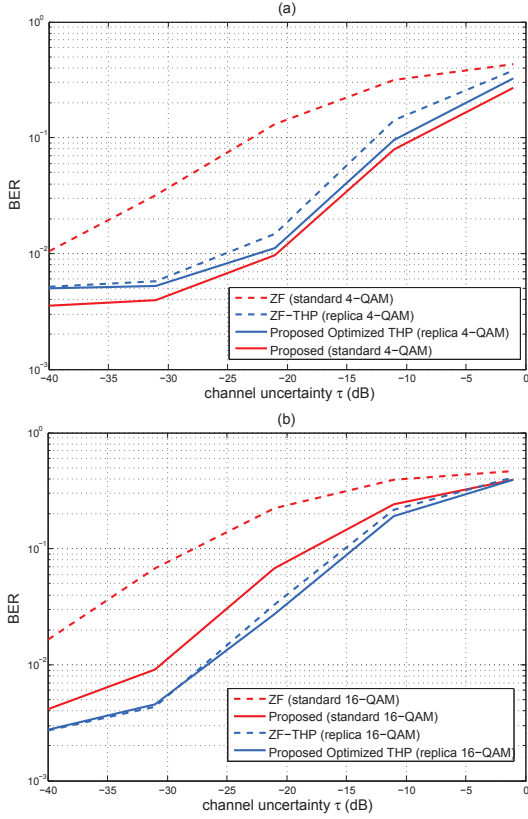


Fig. 11. Uncoded BER versus the normalized channel uncertainty ($K = N = 32$, $L = 8$): (a) standard and replica 4-QAM; (b) standard and replica 16-QAM.

VIII. CONCLUSIONS

In this work, we have introduced a sparse MC-CDMA downlink model and proposed a power-efficient precoding method under minimum symbol error probability (SEP) requirements at the MSs. It has been shown that when the system load is high, the proposed precoder significantly reduces the transmission power over (regularized) zero-forcing based precoders under the same SEP targets. It has also been shown that using the proposed precoder, the sparse MC-CDMA system with a proper choice of sparsity level attains almost the same power efficiency and link throughput as that of the dense MC-CDMA, but with lower complexity. These features, along with the fact that channel measurements may be simplified with sparse signatures, add to the practical appeal of the sparse MC-CDMA and make it a valuable candidate for future-generation wireless communication systems.

APPENDIX A

CRITICAL POINTS ON THE BOUNDARY OF $\mathcal{B}(d_k)$ FOR THE 16-QAM SIGNALING

For the 16-QAM signaling, the SEP target (4) indicates that

$$\underbrace{\frac{1}{\sqrt{2\pi}\sigma} \int_{d_k^{(r)} - \bar{y}_k^{(r)} - \rho_-^{(r)}}^{d_k^{(r)} - \bar{y}_k^{(r)} + \rho_+^{(r)}} e^{-\frac{(z_k^{(r)})^2}{2\sigma^2}} dz_k^{(r)}}_{\mathcal{O}^{(r)}} \geq 1 - Pe_k, \quad (34)$$

TABLE III
INTEGRAL PARAMETERS FOR THE 16-QAM SIGNALING

$\rho_-^{(r)} / \rho_+^{(i)}$	$\rho_-^{(i)} / \rho_+^{(i)}$	$+\infty / \beta_k$	β_k / β_k	$\beta_k / +\infty$
$+\infty / \beta_k$	D_1	$D_{\{2,4\}}$	D_3	
β_k / β_k	$D_{\{5,13\}}$	$D_{\{6,8,14,16\}}$	$D_{\{7,15\}}$	
$\beta_k / +\infty$	D_9	$D_{\{10,12\}}$	D_{11}	

$$\underbrace{\frac{1}{\sqrt{2\pi}\sigma} \int_{d_k^{(i)} - \bar{y}_k^{(i)} - \rho_-^{(i)}}^{d_k^{(i)} - \bar{y}_k^{(i)} + \rho_+^{(i)}} e^{-\frac{(z_k^{(i)})^2}{2\sigma^2}} dz_k^{(i)}}_{\mathcal{O}^{(i)}} \geq 1 - Pe_k, \quad (34)$$

where $\rho_+^{(r)} / \rho_+^{(i)}$ and $\rho_-^{(r)} / \rho_-^{(i)}$ relate to the decision regions and parameterize the upper/lower bounds for the integrals. Table III specifies these parameters for different constellation points.

Given symbol $d_k \in \mathcal{D}_k$ and a target Pe_k , one can determine the precise constraint region $\mathcal{B}(d_k)$ on noiseless output \bar{y}_k at MS k from inequality (34). In particular, the boundary of the region is determined by the equality $\mathcal{O}^{(r)} \mathcal{O}^{(i)} = 1 - Pe_k$ in (34). In the following, we explain how to determine a set of critical points on the boundary of the corresponding regions for three representative constellation points $\{D_{11}, D_{12}, D_{16}\}$.

For the corner constellation point D_{11} , the critical boundary points can be determined according to three different combinations of $(\mathcal{O}^{(r)}, \mathcal{O}^{(i)})$:

- (i) $(1 - Pe_k, 1)$: $\bar{y}_k^{(r)} = 2\beta_k - \sigma Q^{-1}(1 - Pe_k)$, $\bar{y}_k^{(i)} = +\infty$;
- (ii) $(1, 1 - Pe_k)$: $\bar{y}_k^{(r)} = +\infty$, $\bar{y}_k^{(i)} = 2\beta_k - \sigma Q^{-1}(1 - Pe_k)$;
- (iii) $(\sqrt{1 - Pe_k}, \sqrt{1 - Pe_k})$: $\bar{y}_k^{(r)} = \bar{y}_k^{(i)} = 2\beta_k - \sigma Q^{-1}(\sqrt{1 - Pe_k})$,

where $Q^{-1}(\cdot)$ denotes the inverse of the standard Q -function.

For the center constellation point D_{16} , the critical boundary points can be found by examining the following combinations of $(\mathcal{O}^{(r)}, \mathcal{O}^{(i)})$:

- (i) $(\frac{1 - Pe_k}{\alpha}, \alpha)$: $\bar{y}_k^{(r)} = \beta_k \pm \delta_1$, $\bar{y}_k^{(i)} = \beta_k$;
- (ii) $(\alpha, \frac{1 - Pe_k}{\alpha})$: $\bar{y}_k^{(r)} = \beta_k$, $\bar{y}_k^{(i)} = \beta_k \pm \delta_1$;
- (iii) $(\sqrt{1 - Pe_k}, \sqrt{1 - Pe_k})$: $\bar{y}_k^{(r)} = \beta_k \pm \delta_0$, $\bar{y}_k^{(i)} = \beta_k \pm \delta_0$,

where $\alpha = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\beta_k}^{\beta_k} e^{-\frac{v^2}{2\sigma^2}} dv \geq \sqrt{1 - Pe_k}$, parameters δ_0 and δ_1 are chosen to satisfy:

$$\begin{cases} Q\left(\frac{\delta_0 - \beta_k}{\sigma}\right) - Q\left(\frac{\delta_0 + \beta_k}{\sigma}\right) = \sqrt{1 - Pe_k}, \\ Q\left(\frac{\delta_1 - \beta_k}{\sigma}\right) - Q\left(\frac{\delta_1 + \beta_k}{\sigma}\right) = \frac{1 - Pe_k}{\alpha}. \end{cases} \quad (35)$$

For the side constellation point D_{12} , the critical boundary points can be found similarly by examining the following combinations of $(\mathcal{O}^{(r)}, \mathcal{O}^{(i)})$:

- (i) $(\alpha, \frac{1 - Pe_k}{\alpha})$: $\bar{y}_k^{(r)} = 2\beta_k - \sigma Q^{-1}(\frac{1 - Pe_k}{\alpha})$, $\bar{y}_k^{(i)} = \beta_k$;
- (ii) $(\sqrt{1 - Pe_k}, \sqrt{1 - Pe_k})$: $\bar{y}_k^{(r)} = 2\beta_k - \sigma Q^{-1}(\sqrt{1 - Pe_k})$, $\bar{y}_k^{(i)} = \beta_k \pm \delta_0$;
- (iii) $(1, 1 - Pe_k)$: $\bar{y}_k^{(r)} = +\infty$, $\bar{y}_k^{(i)} = \beta_k \pm \delta_2$,

where parameter δ_2 is chosen to satisfy:

$$Q\left(\frac{\delta_2 - \beta_k}{\sigma}\right) - Q\left(\frac{\delta_2 + \beta_k}{\sigma}\right) = 1 - Pe_k. \quad (36)$$

REFERENCES

- [1] N. Yee, J. Linnartz, and G. Fettweis, "Multi-carrier CDMA in indoor wireless radio networks," in *Proc. of IEEE Personal Indoor and Mobile Radio Commun.*, 1993.
- [2] S. Hara and R. Prasad, "Overview of multicarrier CDMA," *IEEE Commun. Mag.*, vol. 35, no. 12, pp. 126–133, 1997.
- [3] A. M. Tulino, L. Li, and S. Verdú, "Spectral efficiency of multicarrier CDMA," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 479–505, 2005.
- [4] K. Fazel and S. Kaiser, *Multi-carrier and spread spectrum systems: from OFDM and MC-CDMA to LTE and WiMAX*. Wiley, 2008.
- [5] M. Yoshida and T. Tanaka, "Analysis of sparsely-spread CDMA via statistical mechanics," in *Proc. IEEE Int. Symp. Inf. Theory*, 2006.
- [6] A. Montanari and D. Tse, "Analysis of belief propagation for non-linear problems: The example of CDMA (or: How to prove Tanaka's formula)," in *Proc. of Inf. Theory and Appl. Workshop*, 2006.
- [7] J. Raymond and D. Saad, "Sparsely spread CDMA—statistical mechanics-based analysis," *J. Phys. A: Math. Theor.*, vol. 40, no. 41, pp. 12 315–12 333, 2007.
- [8] D. Guo and C. Wang, "Multiuser detection of sparsely spread CDMA," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 3, pp. 421–431, 2008.
- [9] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channel," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1616–1626, 2008.
- [10] R. Razavi, M. Al-Imari, M. A. Imran, R. Hoshyar, and D. Chen, "On receiver design for uplink low density signature OFDM (LDS-OFDM)," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3499–3508, 2012.
- [11] J. Choi, "Low density spreading for multicarrier systems," in *IEEE Int. Symp. on Spread Spectrum Tech. and Appl.*, 2004, pp. 575–578.
- [12] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, 2005.
- [13] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part II: Perturbation," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 537–544, 2005.
- [14] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, 2006.
- [15] L. Sanguinetti and M. Morelli, "Non-linear pre-coding for multiple-antenna multi-user downlink transmissions with different QoS requirements," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 852–856, 2007.
- [16] R. Muller, D. Guo, and A. L. Moustakas, "Vector precoding for wireless MIMO systems and its replica analysis," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 3, pp. 530–540, 2008.
- [17] M. B. Shenouda and T. N. Davidson, "Nonlinear and linear broadcasting with QoS requirements: Tractable approaches for bounded channel uncertainties," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1936–1947, 2009.
- [18] J. Kang, H. Ku, D.-S. Kwon, and C. Lee, "Tomlinson-Harashima precoder with tilted constellation for reducing transmission power," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3658–3667, 2009.
- [19] A. Garcia-Rodriguez and C. Masouros, "Power-efficient Tomlinson-Harashima precoding for the downlink of multi-user MISO systems," *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 1884–1896, 2014.
- [20] E. Bjornson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142–148, 2014.
- [21] A. Silva and A. Gameiro, "Pre-filtering antenna array for downlink TDD MC-CDMA systems," in *Proc. of IEEE Veh. Technol. Conf.*, April 2003.
- [22] T. Salzer and D. Mottier, "Downlink strategies using antenna arrays for interference mitigation in multi-carrier CDMA," in *Proc. of MCSS*, September 2003.
- [23] I. Cosovic, S. Sand, and R. Raulefs, "A non-linear precoding technique for downlink MC-CDMA," in *Proc. of IEEE Veh. Technol. Conf.*, May 2005.
- [24] M. Morelli and L. Sanguinetti, "A unified framework for Tomlinson-Harashima precoding in MC-CDMA and OFDMA downlink transmissions," *IEEE Trans. Commun.*, vol. 55, no. 10, pp. 1963–1972, 2007.
- [25] C. Masouros and E. Alsusa, "Soft linear precoding for the downlink of DS/CDMA communication systems," *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 203–215, 2010.
- [26] M. Li, C. Liu, and S. V. Hanly, "Distributed base station cooperation with finite alphabet and QoS constraints," in *Proc. of IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, July 2013.
- [27] C. Liu, M. Li, and S. V. Hanly, "Power-efficient transmission design for MISO broadcast systems with QoS constraints," in *Proc. of Inf. Theory and Appl. Workshop*, San Diego, CA, USA, February 2013.
- [28] M. Li, C. Liu, and S. V. Hanly, "Distributed transmitter optimization for the Wyner-type downlink via belief propagation," *IEEE Commun. Lett.*, vol. 18, no. 3, pp. 471–474, 2014.
- [29] J. Proakis and M. Salehi, *Digital Communications*, 5th ed. McGraw-Hill, 2007.
- [30] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.
- [31] Q. Zheng, Y. Chen, H. Lee, R. Dreslinski, C. Chakrabarti, A. Anas-tasopoulos, S. Mahlke, and T. Mudge, "Using graphics processing units in an LTE base station," *J. Signal Process. Syst.*, 2014, DOI: 10.1007/s11265-014-0932-x.
- [32] China Mobile Research Institute, "C-RAN: The road towards green RAN," 2013, white paper, version 3.0.
- [33] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1439–1451, 2006.
- [34] P. Giselsson, M. D. Doan, T. Keviczky, B. De Schutter, and A. Rantzer, "Accelerated gradient methods and dual decomposition in distributed model predictive control," *Automatica*, vol. 49, no. 3, pp. 828–833, 2013.
- [35] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, "Pre-coding in multiantenna and multiuser communications," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1305–1316, 2004.
- [36] G. D. Golden, C. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electron. Lett.*, vol. 35, no. 1, pp. 14–16, 1999.